Applied Physiology, Nutrition, and Metabolism
Physiologie appliquée, nutrition et métabolisme

Comparison between parameters from maximal cycle ergometer test first without respiratory gas analysis and thereafter with respiratory gas analysis among healthy prepubertal children

SCHOLARONE™
Manuscripts

Comparison between parameters from maximal cycle ergometer test first without respiratory gas analysis and thereafter with respiratory gas analysis among healthy prepubertal children

Tuomo T. Tompuri, Niina Lintu, Sonja Soininen, Tomi Laitinen and Timo Antero Lakka

Corresponding author:

Tuomo Tompuri (email. tuomo.tompuri@kuh.fi; tel. +358447113203; fax. +35817173912)

Affiliations:

T. Tompuri (tuomo.tompuri@kuh.fi) and T. Laitinen (tomi.laitinen@kuh.fi): Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, P.O. Box 100, FIN-70029, Kuopio, Finland.

N. Lintu (niina.lintu@uef.fi), S. Soininen (spkorpel@student.uef.fi) and T. A. Lakka (timo.lakka@uef.fi): Institute of Biomedicine (Physiology), University of Eastern Finland, P.O. Box 1627, FIN-70211, Kuopio, Finland.

Abstract

It is important to distinguish true and clinically relevant changes and methodological noise from measure to measure. In the clinical practice maximal cycle ergometer tests are typically performed first without respiratory gas analysis, and thereafter, if needed, with respiratory gas analysis. Therefore we report comparison of parameters from maximal cycle ergometer exercise tests which were done first without respiratory gas analysis and thereafter with it in 38 prepubertal and healthy children (20 girls, 18 boys). The Bland-Altman method was used to assess agreement in maximal workload (W_{MAX}) and heart rate (HR) and systolic blood pressure (SBP) between rest and maximum. Girls achieved higher W_{MAX} in the exercise tests with respiratory gas analysis compared with exercise tests without respiratory gas analysis (P=0.016), whereas W_{MAX} was similar in the tests among boys. Maximal HR (proportional offset -1%; CV% 3.3) and highest SBP (proportional offset 3%; CV% 10.6) were similar in the tests among children. Precision and agreement for HR improved and precision for SBP worsened with increasing exercise intensity. Heteroscedasticity was not observed for W_{MAX}, HR or SBP. We conclude that maximal cycle ergometer tests without and with respiratory gas analysis can be used consecutively because measurement of respiratory gases did not impair performance or have significant effect on the maximality of the exercise tests. Our results suggest that similar references can be used for children who accept or refuse using a mask during maximal exercise test.

Key words: blood pressure; cycle ergometer; exercise test; heart rate; maximal workload; Wmax

Introduction

Measurement of maximal oxygen uptake (VO_{2MAX}) is a criterion procedure in the evaluation of cardiorespiratory fitness (Fletcher et al. 2001; Wasserman et al. 2005). While VO_{2MAX} refers to the capacity of the human body to increase its energy metabolism during maximal exertion, maximal workload (W_{MAX}) reflects also neuromuscular performance (Nickleberry and Brooks 1996). During incremental dynamic exercise oxygen uptake and performed work depend on each other (Dennis and Noakes 1998), and although VO_{2MAX} limits W_{MAX}, VO_{2MAX} is rather a result than a cause for performed work. Thus cardiorespiratory capacity can also be assessed by W_{MAX} achieved in a cycle ergometer exercise test (Wasserman et al. 2005). Moreover, clinical exercise testing has diagnostic and prognostic value, and it also includes evaluation of the hemodynamic response during exercise (Fletcher et al. 2001; Gläser et al. 2013; Tavel 2001).

Repeatability for measures obtained from maximal exercise tests in children has been rarely reported, although agreement analysis is needed to guarantee valid interpretation of the exercise test findings. There is only one published report on the agreement of measures received from treadmill exercise tests among healthy children (Johnston et al. 2005), and there are no such studies using a cycle ergometer. Moreover, agreement of maximal cycle ergometer tests with respiratory gas analysis (CE^{RGA} test) and without it (CE test) has not previously been reported in children. Regardless of methodological priority of the directly measured VO_{2MAX} (Armstrong and Welsman 2008; Wasserman et al. 2005), simpler test protocol might be sometimes more usable. Furthermore, respiratory gas analysis may not always be available because of lack of technical facilities or because some children might refuse to use a mask, which is needed to measure respiratory gases (Eiberg et al. 2005). Typically in clinical practice, CE tests are performed primarily, and CE^{RGA} tests are performed to attain more detailed data on demand. Therefore scientific evidence is needed to justify consecutive use of pediatric cycle ergometer tests without and with respiratory gas analysis.

We hypothesized that there may be some discrepancy between the results of CE and CE^{RGA} tests. Thus, to facilitate valid interpretation and to identify factual changes between CE and CE^{RGA} tests, it is important to investigate whether respiratory gas analysis itself has an influence on the maximal level of the exercise test or has an effect on achieved maximal workload or hemodynamic response in healthy prepubertal children. We therefore studied the agreement of W_{MAX} and hemodynamic responses to exercise between CE and CE^{RGA} tests to evaluate whether the respiratory gas analysis affects maximal performance and measures during the maximal test among children.

Methods

Subjects

The present data are based on baseline examinations of the Physical Activity and Nutrition in Children (PANIC) Study, which was carried out between October 2007 and December 2009 in 512 children who were 6-8 years of age living in the city of Kuopio. Firstly, the children underwent a maximal exercise test on a cycle ergometer without respiratory gas analysis. For the present study, 40 children aged 7.2-8.3 years underwent another maximal exercise test with respiratory gas analysis 5 to 14 days after the first exercise test. Data from two children were excluded because of technical problems or asthmatic symptoms during the exercise test. Therefore, the present analyses included data from 38 children (20 girls, 18 boys). Of these children, 37 were Caucasian and one was African. A physician performed a medical examination for all of the children before the first exercise test. All of the children were healthy and prepubertal (testicular size ≤ 3 ml in boys, breast development stage M1 in girls) according to Tanner's stages (Marshall and Tanner 1969; Marshall and Tanner 1970). Ethical approval for the PANIC Study was granted by the Research Ethics Committee of Hospital District of Northern Savo, and the study was performed according to the Declaration of Helsinki. All of the children and their parents gave their written informed consent.

Exercise test

Maximal exercise tests were performed using the Ergoselect 200K electromagnetic cycle ergometer (Ergoline, Bitz, Germany), which was controlled by the Cardiosoft electrocardiography software, Version 6.5 (GE-Healthcare Medical systems, Freiburg, Germany). The length of the crank arms was 150 mm, and saddle height was set to a knee angle of 160 degrees when the leg on the pedal was extended (LeMura et al. 2001).

The exercise test protocol (Lintu et al. 2014) included a 5-minute pre-exercise resting period lying in the supine position (pre-rest), a 2.5-minute pre-exercise anticipatory phase while sitting on the ergometer, 3 minutes of warm-up cycling with a workload of 5 W, 1 minute of steady-state cycling with a workload of 20 W, cycling an increasing workload of 1 W per 6 seconds until exhaustion, a 4-minute cool-down cycling with a workload of 5 W, and a 5-minute post-exercise resting period lying in the supine position (post-rest). Transition from warm-up cycling at 5 W to steady state cycling at 20 W lasted approximately 15 seconds. Transition from sitting on the ergometer to lying in the supine position after cycling lasted 20-30 seconds.

The children were asked to keep the cadence stable and within 70–80 revolutions per minute. Exhaustion was defined as the inability to maintain the cadence above 65 revolutions per minute regardless of vigorous verbal exhortation. The exercise test was considered maximal if the reason for terminating the test indicated maximal effort and maximal cardiovascular capacity (Armstrong et al. 1996; Karila et al. 2001). We calculated predicted maximal heart rate using a formula 208-(0.7*age), which is valid in children (Machado and Denadai 2011). Total exercise time ranged between 8 minutes 3 seconds and 14 minutes 57 seconds. Maximal workload was defined as the workload at the end of the exercise test to accuracy of 1 W and was expressed not only as absolute but also as weight-proportional and lean mass (LM) -proportional values (Tompuri et al. 2014).

CE tests were performed at 9-12 A.M. and CE^{RGA} tests at 1-4 P.M. In CE^{RGA} tests, respiratory gases were collected using the Jaeger Oxycon Pro gas analyzer (Viasys Healthcare, Hoechberg, Germany) and face masks for children (Hans–Rudolph, Shawnee, Kansas, USA). The

respiratory gases were measured directly using the breath–by–breath method between the anticipatory phase and the cool-down phase of the CE^{RGA} tests. Data on respiratory gas analysis i.e. peak ventilation volume, peak oxygen pulse, peak respiratory exchange ratio, peak ventilatory equivalent for oxygen, ventilatory equivalent for oxygen at ventilatory threshold (Milani et al. 2006), the lowest ventilatory equivalent for carbon dioxide, and the percentage of oxygen uptake at ventilatory threshold vs. maximal oxygen uptake were used to characterize participants and to evaluate if clinically relevant measurements of respiratory gases may be associated with measurement bias. Moreover, to characterize level of cardiorespiratory fitness of the participants in the current study, we compared the maximal workload to population reference (Lintu et al. 2014).

Hemodynamic measurements

Systolic blood pressure (SBP) was measured using an aneroid sphygmomanometer (Heine Gamma G7, Herrsching, Germany). Heart rate (HR) was recorded continuously from pre-rest to post-rest using the Cardiosoft system. SBP and HR used in the analyses were measured 1) at the end of the pre-rest, 2) while sitting on the ergometer before the test, 3) 15 seconds before the end of the warm-up, 4) as close as possible to exhaustion (highest measured SBP and highest recorded HR), 5) soon after exhaustion (SBP immediately after exhaustion and HR 30 seconds after exhaustion), 6) two minutes after exhaustion, 7) four minutes after exhaustion and 8) at the end of the post-rest. During incremental cycling SBP was measured at 2-minute intervals. SBP or HR response refers change from pre-rest to highest SBP.

Body size and composition

LM and body fat percentage were assessed with the Lunar dual–energy X–ray absorptiometry (DXA) device (Lunar Prodigy Advance; GE Medical Systems, Madison, Wisconsin, USA) in the afternoons and in a non–fasting state using manufacturer's instructions (Tompuri et al. 2015). Most of the children (76 %) underwent DXA within the same week as the CE^{RGA} tests, and the time gap between CE^{RGA} tests and DXA was up to 22 days. Body height was measured using a wall–mounted stadiometer in Frankfurt plane without shoes to an accuracy of 0.1 cm. Body weight was

measured with the InBody$_{720}$ device (Biospace CO. Ltd, Seoul, Korea) to an accuracy of 0.1 kg after an overnight fast on the same day as the CE tests. Standard deviation scores (SDS) for BMI were calculated using age- and sex-specific reference data (Saari et al. 2009) to characterize the study sample. Respectively, the prevalence of overweight and obesity was assessed using the age- and sex-specific BMI cut-offs of the International Obesity Task Force (Cole et al. 2000).

Statistical analyses

Statistical analyses were performed with the IBM SPSS Statistics for Windows, Version 19.0 (IBM Corp., Armonk, NY, USA), and significance was set at *P<0.05*. The Shapiro-Wilk test was used to test distribution normality: most distributions were parametric, and nonparametric distributions have been indicated as NP in the results. Pearson's coefficients for correlation were used to assess the magnitude of the associations between parameters obtained from the CE and CE^{RGA} tests. The paired T–test was used for normally distributed variables or the Wilcoxon signed rank test was used for variables with skewed distributions to test the significance of the differences between the parameters obtained from the CE and CE^{RGA} tests.

Agreement analyses included determination of bias, limits of agreement (LA), heteroscedasticity, coefficients of variation (CV%) and intra–class correlation coefficients (ICC). Bias refers to the difference between values from the CE and CE^{RGA} tests (*value in CE^{RGA} test - value in CE test).* Moreover, Pearson's coefficients for correlation were calculated to explore whether body fat percent or measurements of respiratory gases were associated with bias. Bland–Altman analyses (Bland and Altman 1986) including plot figures were performed to demonstrate agreements between the CE and CE^{RGA} tests. In the plot figures, averages of values from the CE and CE^{RGA} tests were located at the y-axis and bias was located at the x-axis. While LA (mean bias ±2 SD of bias) and CV% (SD of bias divided by mean * 100) indicate precision, bias show trueness. In other words, CV% informs about consistency between the tests, and CV% below 10% indicates acceptable stability of the differences (de Groot et al. 2011). Bias and LAs have been presented in absolute and relative values (%bias and %LA)

because relative values enable a better comparison between the tests. Moreover, heteroscedasticity was determined between bias and mean using Pearson's correlation coefficients for parametric measures and Spearman's correlation coefficients for nonparametric measures. Heteroscedasticity refers to aberration of the bias, which is associated with magnitude of the bias. ICC indicates relative reliability between the tests assessing intrapersonal between measure associations. ICC was computed for single measures using a two-way random model and consistency definition with corresponding 95% confidence intervals (CI).

The independent samples T-test for normally distributed variables and the Mann-Whitney U-test for variables with skewed distribution were used to assess the similarity of bias between sexes. While hemodynamic bias and measures were similar in girls and boys, we performed the hemodynamic agreement analyses only in the entire study sample of children.

Results

Characteristics

Characteristics of children are presented in Table 1. BMI-SDS did not differ statistically significantly from the national reference. One girl and three boys were classified to be overweight or obese (BMI-SDS >1.16 in girls and >0.78 in boys). Moreover, body fat percent, peak ventilation, peak oxygen pulse, ventilatory equivalent for oxygen at maximal exercise or at ventilatory threshold and the lowest value for ventilatory equivalent for carbon dioxide did not correlate statistically significantly with measurement bias. Maximal workload per body weight (W/kg) was in girls similar to population reference but higher in boys (P<0.001).

Relationships between tests

Strong positive correlations were found between the CE and CE^{RGA} tests for W_{MAX} in girls and boys (r ≥0.85; P<0.001), HR measures (r =0.48-0.67; P≤0.002) and SBP measures (r =0.36-0.62; P≤0.026) at different phases of the test. The SBP response to exercise was the only measure that did not correlate statistically significantly between the CE and CE^{RGA} tests (r=0.20, NS). Moreover, heteroscedasticity was not observed in W_{MAX} (Figure 1), HR measures or SBP measures.

Maximal workload

Girls achieved a significantly higher W_{MAX} (+3.4 %) in the CE^{RGA} test than in the CE test (Table 2, Figure 1), whereas there was no significant difference between the tests in boys. However, W_{MAX} bias between the CE and CE^{RGA} tests were similar in girls and boys. LAs for W_{MAX} were narrow and CV%s were low (Table 2, Figure 1). %LAs of the W_{MAX} averaged from -8 % to 15 % in girls and from -9 % to 13 % in boys. Moreover, ICCs for W_{MAX} were statistically significant between the CE and CE^{RGA} tests in girls and in boys (Table 2, Figure 1).

Heart rate

HR values were similar between the CE and CE^{RGA} tests except for HR at rest before exercise (Table 3). At rest before exercise HR was significantly lower in the CE^{RGA} tests as compared with the CE tests. Bias, LAs and CV%s for HR were at appropriate levels after pre-rest and were the best in maximal exercise. ICCs in HR measures were significant (Table 3), but ICCs were lower as compared with W_{MAX} ICC. The mean (SD) percentage of the maximal HR from the predicted maximal HR was 97.4% (3.8) in the CE^{RGA} tests and 98.0 % (4.1) in the CE tests, which refers to a 3.3 CV%, LA from -7.0 % to +5.9 % and high ICC (0.91: 0.83 - 0.95).

Systolic blood pressure

SBP response to exercise and SBP values immediately after maximum were statistically significantly higher in the CE^{RGA} tests compared with the CE tests (Table 4). Correspondingly, bias was quite small except for SBP response and SBP value after maximum. However, during exercise, measured CV% and LA values were worse compared with those of the HR measures (Tables 3 and 4). ICCs of SBP values were statistically significant (Table 4) except for the SBP response to exercise. However, only SBP at recovery 2-minutes after maximum had ICC over 0.60.

Discussion

Our study shows that maximal HR and W_{MAX} had appropriate methodological agreement between maximal cycle ergometer tests, which were performed first without and thereafter with

respiratory gas analysis in prepubertal children. These results suggest that measurement of respiratory gases using a mask did not impair W_{MAX} and had no marked effect on the maximality of exercise in cycle ergometer exercise test among prepubertal children. Correspondingly, during exercise measured SBP agreed poorly indicating weak repeatability. In clinical practice and in scientific research, these results help distinguish between true and clinically significant changes and methodological noise from measure to measure.

Based upon the LA of weight-proportional W_{MAX} in this study, a clinically significant improvement in W_{MAX} refers to a greater than 0.4 W/kg increase between these two tests. Corresponding values for a clinically significant decline were greater than 0.2 W/kg in girls and greater than 0.3 in boys. As compared with the recently published reference values for measures of cardiorespiratory fitness (Lintu et al. 2014), observed bias should not significantly affect fitness evaluation. Moreover, extreme variation as indicated by LA should not markedly alter fitness categories, which suggests that diagnostic confounding is unlikely.

The observed CV% for W_{MAX} (5-6 %) was not only acceptable (<10 %) but also similar to that reported for VO_{2MAX} from treadmill exercise test among children (5-11 %) (Figueroa-Colon et al. 2000; Johnston et al. 2005) or among trained adults (3-8 %) (Bagger et al. 2003; Lourenço et al. 2011; Midgley et al. 2007) or for VO_{2MAX} from cycle ergometer exercise test among cardiac patients (4 %) (Meyer et al. 1997). It is remarkable that while maximal exercise tests were performed without and with respiratory gas analysis, maximal performance consistency was similar (5-6 %) to a situation in which both tests were performed with respiratory gas analysis (5 %) (Johnston et al. 2005).

There is no explanation why respiratory gas analysis would improve the achieved W_{MAX} that was observed in our study among girls. Previously significant improvement in VO_{2MAX} has not been found in repeated tests, suggesting that learning effects do not contribute to results among children while using a methodological combination of a treadmill and respiratory gas analysis (Eiberg et al. 2005; Johnston et al. 2005). On the other hand, the currently observed trend towards a better maximal

performance (mean difference: 2-3 ±6%) was similar to earlier observations among children (3 ±5%) and untrained adults (2 ±6%) who used a treadmill and respiratory gas analysis in both tests (Johnston et al. 2005). Therefore, we can assume that respiratory gas analysis per se would unlikely affect performance in the maximal exercise test. However, an explanation for the significant improvement in girls remains obscure. In the earlier studies (Eiberg et al. 2005; Johnston et al. 2005), analyses of smaller cohorts have been performed together for both sexes, and potential differences in familiarization with the exercise test laboratory and the experience of maximal exercise between sexes (Allen et al. 1993; Kajantie and Philips 2006; Merz et al. 2010) are rather speculative.

HR can be used to assess the intensity level of exercise (Achten and Jeukendrup 2003) and to explore cardiac response to exercise in more detail (Savonen et al. 2006). We did assess the maximality of exercise tests similarly to clinical practice, and marginal average bias and low CV% for maximal HR were observed similarly to previous studies (Figueroa-Colon et al. 2000; Johnston et al. 2005). Currently scored average of the maximal HR was even higher or similar to other pediatric studies (Armstrong et al. 1996; Eiberg et al. 2005; Figueroa-Colon et al. 2000; Johnston et al. 2005). It is relevant that repeated testing may reveal relevant changes in maximal HR even if the maximal HR would be within population reference (Lintu et al. 2014) in both exercise tests.

While HR had best agreement at maximal exercise, HR measurements during sub-maximal exercise, i.e., during warming up or cooling down periods showed acceptable agreement including consistency and predominantly good reliability. In general, HR measure biases were minimal, demonstrating satisfying trueness. Thus, methodology should not confound results if performing comparisons of groups using HR averages. However, precision, consistency and reliability got slightly worse at lower exercise levels and lower heart rates. At rest, especially at pre-rest, HR agreement was poor. This was reflected by weak reliability as indicated by low ICC (<0.60) and poor consistency as suggested by high CV% (>10%). This observed trend is consistent with physiological principles, i.e., HR varies from beat-to-beat more at rest because of stronger vagal autonomic tone (Coote 2010; Task Force

of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology 1996). However, methodological HR measurement should be specific because of ECG recording.

SBP measures generally had poor agreement during exercise. Although a predominantly small bias indicated acceptable trueness, SBP measures had a weak level of precision as suggested by wide LAs, poor consistency as illustrated by high CV%s (>10 %), and weak reliability as demonstrated by low ICCs (<0.60). While also corresponding inter-individual references are wide (Lintu et al. 2014), our results reflect a vast spectrum of intra-individual measures and suggest that SBP measures during exercise would have limited diagnostic value for clinical evaluation. Our results suggest that greater than 32 mmHg i.e. 25 % increase in maximal SBP would be considered to be real change irrespective of methodological noise. While during exercise measured SBP has been used as a valuable diagnostic and predictive indicator among adults (Tavel 2001), further studies assessing SBP measure agreement during exercise should be performed among pediatric patients to confirm our results.

This is the first study on the agreement between two maximal cycle ergometer exercise tests in children. We had larger study sample compared with earlier studies assessing agreement or relationships between exercise tests, which were performed by treadmill (Eiberg et al. 2005; Johnston et al. 2005). Moreover, our cohort represented a uniform prepubertal age group, and maturity related variation should not confound the current results. We also performed agreement analysis for SBP measures, which have not been published previously. Furthermore, the study scheme including order of exercise tests was planned to serve clinical practice. Circadian variation should be taken into account as a potential limitation of the study because the CE test was performed before noon and the CE^{RGA} test was performed in the afternoon. HR generally tends to be lowest in the morning and at highest in late afternoon (Drust et al. 2005), but circadian rhythms should not affect the HR (Cohen 1980) hemodynamic response in maximal exercise (Reilly et al. 2000) and neither cycling performance (Dalton 1997). We observed significant differences in pre-rest HR levels, but the observed difference was the opposite compared with that presented by Drust et al. (2005).

We conclude that maximal cycle ergometer exercise tests, which were performed first without and thereafter with respiratory gas analysis, had an acceptable agreement for W_{MAX} and for maximal HR, which justifies consecutive use of these tests in the clinical practice and enables use of similar references for those children who use or might refuse to use a mask during exercise testing. However, SBP had poor agreement during exercise. Agreement of the given values enables us to distinguish factual changes from measurement related noise. Furthermore, repeated tests and clinically relevant changes may possess better diagnostic potential for W_{MAX} and for maximal HR compared with population references.

Acknowledgments

Conflict of interest and ethical statements

The authors declare that they have no conflict of interest. These results have not been published elsewhere. The authors declare that these experiments comply with current Finnish laws. The study protocol was approved by the Research Ethics Committee of the Hospital District of Northern Savo. All of the children and their parents gave their informed written consent. The study was performed in accordance with the ethical standards of the Declaration of Helsinki.

References

Achten, J., and Jeukendrup, A.E. 2003. Heart rate monitoring: applications and limitations. Sports Med. **33**(7):517-38.

Allen, M.T., Stoney, C.M., Owens, J.F., Matthews, K.A. 1993. Hemodynamic adjustments to laboratory stress: The influence of gender and personality. Psychosomatic Medicine. **55**(6):505-17. doi: 10.1097/00006842-199311000-00006

Armstrong, N., and Welsman, J.R. 2008. Aerobic fitness. *In* Paediatric Exercise Science and Medicine. *Edited by* N. Armstrong and W. van Mechelen. Oxford University Press, Oxford. pp. 97-108.

Armstrong, N., Welsman, J., and Winsley, R. 1996. Is Peak VO2 a maximal index of children's aerobic fitness? Int. J. Sports Med. **17**(5):356-9. doi: 10.1055/s-2007-972860

Bagger, M., Petersen, P.H., and Pedersen, P.K. 2003. Biological variation in variables associated with exercise training. Int. J. Sports Med. **24**(6):433-40. doi: 10.1055/s-2003-41180

Bland, J.M., and Altman, D.G. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. **327**(8476):307-10. doi: 10.1016/S0140-6736(86)90837-8

Cohen, C.J. 1980. Human circadian rhythms in heart rate response to a maximal exercise stress. Ergonomics. **23**(6):591-5. doi: 10.1080/00140138008924773

Cole, T.J., Bellizzi, M.C., Flegal, K.M., and Dietz, W.H. 2000. Establishing a standard definition for child overweight and obesity worldwide: international survey. BMJ. **320**(7244):1240–3. doi: 10.1136/bmj.320.7244.1240

Coote, J.H. 2010. Recovery of heart rate following intense dynamic exercise. Exp. Physiol. **95**(3):431-40.

Dalton, D., McNaughton, L., and Davoren, B. 1997. Circadian rhythms have no effect on cycling performance. Int. J. Sports Med. **18**(7):538-42. doi: 10.1055/s-2007-972678

de Groot, J.F., Takken, T., Gooskens, R.H., Schoenmakers, M.A., Wubbels, M., Vanhees, L., et al. 2011. Reproducibility of maximal and submaximal exercise testing in "normal ambulatory" and "community

ambulatory" children and adolescents with spina bifida: which is best for the evaluation and application of exercise training? Phys. Ther. **91**(2):267-76. doi: 10.2522/ptj.20100069

Dennis, S.C., and Noakes, T.D. 1998. Physiological and metabolic responses to increasing work rates: relevance for exercise prescription. J. Sports Sci. **16**(Suppl):S77-84. doi: 10.1080/026404198366704

Drust, B., Waterhouse, J., Atkinson, G., Edwards, B., and Reilly, T. 2005. Circadian rhythms in sports performance – an update. Chronobiology International. **22**(1):21-44. doi: 10.1081/CBI-200041039

Eiberg, S., Hasselstrom, H., Grønfeldt, V., Froberg, K., Svensson, J., and Andersen, L.B. 2005. Maximum oxygen uptake and objectively measured physical activity in Danish children 6–7 years of age: the Copenhagen school child intervention study. Br. J. Sports Med. **39**(10):725–30. doi: 10.1136/bjsm.2004.015230

Figueroa-Colon, R., Hunter, G.R., Mayo, M.S., Aldridge, R.A., Goran, M.I., and Weinsier, R.L. 2000. Reliability of treadmill measures and criteria to determine VO2max in prepubertal girls. Med. Sci. Sports Exerc. **32**(4):865-9.

Fletcher, G.F., Balady, G.J., Amsterdam, E.A., Chaitman, B., Eckel, R., Fleg, J., et al. 2001. Exercise standards for testing and training: a statement for healthcare professionals from the American Heart Association. Circulation. **104**(14):1694-740. doi: 10.1161/hc3901.095960

Gläser, S., Friedrich, N., Koch, B., Schäper, C., Völzke, H., Felix, S.B., et al. 2013. Exercise blood pressure and heart rate reference values. Heart Lung Circ. **22**(8):661-7. doi: 10.1016/j.hlc.2013.02.007

Johnston, K.N., Jenkins, S.C., and Stick, S.M. 2005. Repeatability of peak oxygen uptake in children who are healthy. Pediatr. Phys. Ther. **17**(1):11-7. doi: 10.1097/01.PEP.0000154340.43505.91

Kajantie, E., and Phillips, D.I.W. 2006. The effects of sex and hormonal status on the physiological response to acute psychosocial stress. Psychoneuroendocrinology. **31**(2):151–78. doi: 10.1016/j.psyneuen.2005.07.002

Karila, C., de Blic, J., Waernessyckle, S., Benoist, M.R., and Scheinmann, P. 2001. Cardiopulmonary exercise testing in children: an individualized protocol for workload increase. Chest. **120**(1):81-7.

LeMura, L.M., von Duvillard, S.P., Cohen, S.L., Root, C.J., Chelland, S.A., Andreacci, J., et al. 2001. Treadmill and cycle ergometry testing in 5- to 6-year-old children. Eur. J. Appl. Phys. **85**(5):472-8. doi: 10.1007/s004210100461

Lintu, N., Tompuri, T., Viitasalo, A., Soininen, S., Laitinen, T., Savonen, K., et al. 2014. Cardiovascular fitness and haemodynamic responses to maximal cycle ergometer test in children 6-8 years of age. J. Sports Sci. **32**(7):652-9. doi: 10.1080/02640414.2013.845681.

Lourenço, T.F., Martins, L.E., Tessutti, L.S., Brenzikofer, R., and Macedo, D.V. 2011. Reproducibility of an incremental treadmill VO(2)max test with gas exchange analysis for runners. J. Strength Cond. Res. **25**(7):1994-9. doi: 10.1519/JSC.0b013e3181e501d6

Machado, F.A., and Denadai, B.S. 2011. Validity of maximum heart rate prediction equations for children and adolescents. Arq. Bras. Cardiol. **97**(2):136-40.

Marshall, W.A., and Tanner, J.M. Variations in pattern of pubertal changes in girls. 1969. Arch. Dis. Child. **44**(235):291-303. doi:10.1136/adc.44.235.291

Marshall, W.A., and Tanner, J.M. 1970. Variations in pattern of pubertal changes in boys. Arch. Dis. Child. **45**(239):13-23. doi:10.1136/adc.45.239.13

Merz, C.J., Tabbert, K., Schweckendiek, J., Klucken, T., Vaitl,, D., Stark, R., et al. 2010. Investigating the impact of sex and cortisol on implicit fear conditioning with fMRI. Psychoneuroendocrinology. **35**(1):33-46. doi: 10.1016/j.psyneuen.2009.07.009

Meyer, K., Westbrook, S., Schwaibold, M., Hajric, R., Peters, K., and Roskamm, H. 1997. Short-term reproducibility of cardiopulmonary measurements during exercise testing in patients with severe chronic heart failure. Am. Heart J. **134**(1):20-6. doi: 10.1016/S0002-8703(97)70102-9

Midgley, A.W., McNaughton, L.R., and Carroll, S. 2007. Effect of the VO2 time-averaging interval on the reproducibility of VO2max in healthy athletic subjects. Clin. Physiol. Funct. Imaging. **27**(2):122-5. doi: 10.1111/j.1475-097X.2007.00725.x

Milani, R.V., Lavie, C.J., Mehra, M.R., and Ventura, H.O. 2006. Understanding the basics of cardiopulmonary exercise testing. Mayo Clin. Proc. **81**(12):1603-11. doi: 10.4065/81.12.1603

Nickleberry, B.L., Jr, and Brooks, G.A. 1996. No effect of cycling experience on leg cycle ergometer efficiency. Med. Sci. Sports Exerc. **28**(11):1396-1401. doi: 10.1097/00005768-199611000-00008

Reilly, T., Atkinson, G., and Waterhouse, J. 2000. Chronobiology and physical performance. *In* Exercise and sport science. *Edited by* W.E. Garrett Jr and D.T. Kirkendall. Lippincott Williams and Wilkins, Philadelphia, PA. pp. 351-72.

Saari, A., Sankilampi, U., Hannila, M.L., Kiviniemi, V., Kesseli, K., and Dunkel, L. 2011. New Finnish growth references for children and adolescents aged 0 to 20 years: Length/height–for–age, weight–for–length/height, and body mass index–for–age. Ann. Med. **43**(3):235–48. doi: 10.3109/07853890.2010.515603

Savonen, K.P., Lakka, T.A., Laukkanen, J.A., Halonen, P.M., Rauramaa, T.H., Salonen, J.T., et al. 2006. Heart rate response during exercise test and cardiovascular mortality in middle-aged men. Eur. Heart J. **27**(5):582-8. doi: 10.1093/eurheartj/ehi708

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. 1996. Heart rate variability standards of measurement, physiological interpretation, and clinical use. Eur. Heart J. **17**(3):354-381.

Tavel, M.E. 2001. Stress testing in cardiac evaluation: current concepts with emphasis on the ECG. Chest. **119**(3):907-25. doi: 10.1378/chest.119.3.907

Tompuri, T.T., Lakka, T.A., Hakulinen, M., Lindi, V., Laaksonen, D.E., Kilpeläinen, T.O., et al. 2015. Assessment of body composition by dual-energy X-ray absorptiometry, bioimpedance analysis and anthropometrics in children: the Physical Activity and Nutrition in Children study. Clin. Physiol. Funct. Imaging. **35**(1):21-33. doi: 10.1111/cpf.12118.

Tompuri, T., Lintu, N., Savonen, K., Laitinen, T., Laaksonen, D., Jääskeläinen, J., et al. 2015. Measures of cardiorespiratory fitness in relation to measures of body size and composition among children. Clin. Physiol. Funct. Imaging. 35(6):469-77. doi: 10.1111/cpf.12185.

Wasserman, K., Hansen, J.E., Sue, D.Y., Stringer, W.W., Whipp, B.J. 2005. Principles of Exercise Testing and Interpretation Including Pathophysiology and Clinical Applications. Lippincott Williams & Wilkins, Philadelphia, PA.

TABLE 1   Characteristics of children

	Girls (n = 20)		Boys (n = 18)	
	Mean	SD	Mean	SD
Age [year]	7.8	0.3	7.7	0.3
Body height [cm]	129.5	5.5	130.3	4.4
Body weight [kg]	26.7	3.5	27.4	3.4
Body mass index - standard deviation score	-0.24	0.89	-0.09	0.80
Body fat percent [%]	22.1	6.5	15.3	6.7
Lean body mass [kg]	19.6	1.9	22.4	2.0
VO_{2PEAK} [ml·min^{-1}]	1207	132	1510	197
VO_{2PEAK}·kg of body weight^{-1} [ml·min^{-1}·kg^{-1}]	45.8	5.4	55.5	7.0
VO_{2PEAK}·kg of lean mass^{-1} [ml·min^{-1}·kg^{-1}]	61.7	6.2	67.4	5.2
Peak respiratory exchange ratio	1.08	0.07	1.07	0.05
Peak O_2 pulse [ml·beat^{-1}]	6.1	0.6	7.6	1.0
Peak O_2 pulse·kg of lean mass^{-1} [ml·beat·kg^{-1}]	0.31	0.03	0.34	0.02
Peak ventilation [l·min^{-1}]	49	9	61	6
Peak $VE·VO_2^{-1}$	37	5	39	4
$VE·VO_2^{-1}$ at ventilatory threshold	25	2	26	2
Oxygen uptake at ventilatory threshold vs. VO_{2PEAK} [%]	58	5	57	4
The lowest $VE·VCO_2^{-1}$	28	2	28	2

Standard deviation score (SDS) was calculated according to national reference (Saari et al. 2011); Lean body mass and body fat percent were analyzed by dual-energy X-ray absorptiometry; VO_{2PEAK} = maximal oxygen uptake; $VE·VO_2^{-1}$ = Ventilatory equivalent for oxygen; $VE·VCO_2^{-1}$ = Ventilatory equivalent for carbon dioxide

TABLE 2 Agreement of maximal workload in maximal exercise test without respiratory gas analysis (CE) and with it (CE^{RGA}) among 20 girls and 18 boys

Maximal workload (W)		Mean (SD)		Bias (LA)	ICC (95% CI)	Dif. (P)	CV%
		CE	CE^{RGA}				
Per lean body mass (kg)	Girls	3.8 (0.4)	3.9 (0.4)	0.1 (-0.3 – 0.6)	0.87**(.69–.94)	**0.013***	5.7
	Boys	4.3 (0.59)	4.3 (0.4)	0.1 (-0.4 – 0.5)	0.87**(.68–.95)	0.201	5.3
Per body weight (kg)	Girls	2.8 (0.4)	2.9 (0.4)	0.1 (-0.2 – 0.4)	0.93**(.82–.97)	**0.015***	5.4
	Boys	3.5 (0.5)	3.6 (0.5)	0.1 (-0.3 – 0.4)	0.88**(.71–.95)	0.381^{NP}	5.3
Absolute	Girls	74 (9)	77 (11)	3 (-6 – 12)	0.90**(.77–.96)	**0.016***	6.1
	Boys	96 (16)	97 (14)	1 (-8 – 11)	0.95**(.86–.98)	0.229	5.1

*Abbreviations: Bias = result in CE^{RGA} test minus result in CE test; CE = maximal cycle ergometer test; CE^{RGA} = maximal cycle ergometer test with respiratory gas analysis; 95% CI = 95% Confidence Interval; CV% = Coefficient of Variation between CE and CE^{RGA} (standard deviation of the offset / mean *100); Dif. (P) = P-value for differences with paired samples T-test; ICC = Intra-class correlation for single measures using two-way random model and consistency definition; LA = Limits of agreement; SD = Standard deviation; W_{MAX} = Maximal workload; *= P-level: <0.05; ** = P-level <0.001;*

TABLE 3   Agreement of heart rate in maximal exercise test without respiratory gas analysis (CE) and with it (CE^{RGA}) in 38 children

Phase	Mean in CE (beats/min)	Mean in CE^{RGA} (beats/min)	Bias (LA) (beats/min)	%Bias (LA) (beats/min)	ICC (95% CI)	Dif. (P)	CV%
Pre-rest	70 (10)	67 (7)	-3 (-21–15)	-4 (-28–20)	.46* (.16–.68)	.036^{NP}	13.0
Anticipatory	87 (10)	86 (9)	-1 (-20–18)	-1 (-22–20)	.54** (.27–.73)	.501	10.8
WU	116 (11)	115 (11)	-1 (-22–20)	-1 (-19–17)	.56** (.30–.75)	.511^{NP}	9.0
Maximal	199 (8)	197 (8)	-1 (-14–12)	-1 (-7–6)	.67** (.45–.82)	.261	3.3
CD+30”	178 (12)	179 (13)	1 (-20–23)	1 (-12–13)	.65** (.42–.80)	.431	5.9
CD+2’	138 (12)	136 (13)	-2 (-24–20)	-2 (-19–15)	.61** (.36–.78)	.276	8.1
CD+4’	127 (10)	127 (11)	0 (-17–17)	0 (-13–13)	.68** (.47–.82)	.939	6.6
Post-rest	98 (9)	98 (8)	0 (-17–17)	0 (-17–18)	.50** (.23–.71)	.880	8.8
Response	128 (13)	130 (9)	2 (-16–20)	2 (-13–16)	.67** (.45–.82)	.177	6.9

*Anticipatory = sitting on the ergometer before test; Bias = result in CE^{RGA} test minus result in CE test; CD = cooling down, and specific time in seconds (“) or in minutes (‘) indicates time after exhaustion; CE = cycle ergometer test; CE^{RGA} = Cycle ergometer test with respiratory gas analysis; 95% CI = 95% Confidence Interval; CV% = Coefficient of Variation between CE and CE^{RGA} (standard deviation of the offset / mean *100); Dif. (P) = P-value for differences with paired samples T-test for parametric measures, and with Wilcoxon Signed Rank test for nonparametric (NP) measures; ICC = Intra-class correlation for single measures using two-way random model and consistency definition; LA = Limits of agreement; %Bias = Percentage difference between results (bias / mean of the results * 100); Response = from anticipatory phase to maximum;Rest phases were in supine position; SD = Standard deviation; WU = warm-up; *= P-level: <0.05; ** = P-level <0.001*

TABLE 4   Agreement of systolic blood pressure in maximal exercise test with respiratory gas analysis   (CE) and with it (CE^{RGA}) in 38 children

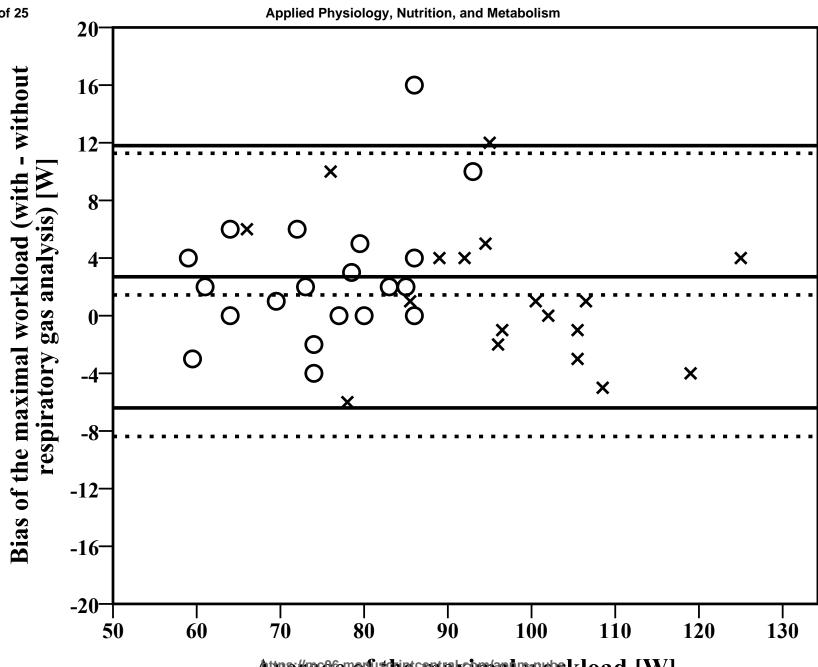| Phase	Mean in CE (mmHg)	Mean in CE^{RGA} (mmHg)	Bias (LA) (mmHg)	%Bias (LA) (mmHg)	ICC (95% CI)	Dif. (P)	CV%
Pre-rest	102 (7)	100 (8)	-2 (-17 – 12)	-2 (-16 – 12)	.50*(.21-.71)	0.133^{NP}	7.2
Anticipatory	103 (8)	100 (7)	-3 (-20 – 15)	-3 (-20 – 15)	.36*(.05-.60)	0.065	8.7
WU	108 (11)	107 (11)	-1 (-23 – 21)	-1 (-22 – 20)	.49*(.20-.70)	0.623	10.3
Highest	129 (13)	133 (12)	4 (-24 – 32)	3 (-19 – 25)	.38*(.07-.62)	0.067	10.6
CD+30"	115 (15)	126 (14)	11 (-21 – 43)	9 (-18 – 37)	.43*(.13-.66)	<0.001	13.2
CD+2'	118 (13)	121 (13)	3 (-20 – 26)	3 (-17 – 23)	.62**(.37-.78)	0.087	9.6
CD+4'	112 (13)	112 (11)	0 (-24 – 23)	0 (-21 – 20)	.55**(.28-.74)	0.843	10.3
Post-rest	107 (7)	105 (7)	-2 (-16 – 13)	-1 (-15 – 12)	.51*(.23-.72)	0.277^{NP}	6.9
Response	26 (13)	33 (11)	7 (-23 – 37)	33 (-126 – 193)	.23 (-.09-.51)	0.007	51.3

Number of the subjects was otherwise 38, but exceptions were pre- and post-rests and CD+4' (n = 37), and WU (n = 36). Abbreviations are by others as presented in table 3.

Figure captions:

FIGURE 1

Bland-Altman plots demonstrating absolute offsets and limits of agreement for maximal workload [W] in the

exercise test with and without respiratory gas analysis for girls (O —) and boys (X ....)